信任博弈范式真的能测量信任吗？争议与展望*

贡喆 [1] 唐玉洁 [1] 刘昌 [2]

([1] 四川师范大学心理学院，成都 610068)

([2] 南京师范大学心理学院，南京 210097)

摘要 信任博弈范式测量信任的有效性一直存在争议。一方面，社会偏好与风险偏好可能在范式变动中对投资信任水平产生影响，降低信任博弈的内部效度。另一方面，信任博弈与量表信任的相关水平较低，这可能和态度与行为差异、测量类型差异以及测量范式的不足有关。虽然存在争议，但总体而言，信任博弈范式仍然是较为适宜的信任测量手段。未来信任博弈范式可在：验证信任博弈范式变动的科学性；进一步澄清信任博弈范式与量表信任的低相关问题；拓展信任博弈范式的测量维度以及提升信任博弈范式的生态效度等方面继续完善。

关键词 信任; 信任博弈; 社会偏好; 风险偏好

1 引言

信任(trust)是建立在对他人的意向或行为的积极预期基础上而敢于托付(愿意承受风险)的一种心理状态(张宁等人, 2011)。长期以来，信任测量都是信任研究中的重点和难点问题，研究者对如何妥善准确地测量信任难以达成统一意见(Brülhart & Usunier, 2012; Corgnet et al., 2016; Gylfason & Olafsdottir, 2017)。依据 Bauer 和 Freitag(2018)的观点，信任测量分为直接测量和间接测量。直接测量即被试对某维度信任在李克特量表上直接报告，如人际信任量表(interpersonal trust scales, ITS)(Rotter, 1967)；间接测量则通过个体的认知决策、具体行为等推测其信任水平。在信任的间接测量方法中，最为著名的是 Berg 等人(1995)基于 Deutsch(1958)的囚徒困境博弈开发的信任博弈范式(trust game)，以下简称为信任博弈。信任博弈包括信任者(trustor)和被信任者(trustee)两个角色。在经典信任博弈中，博弈双方被分别赋予一定量的真实或虚拟货币，信任者可向被信任者投资任意数量金额，被信任者则获得两倍或三倍不等的增值；接下来，被信任者可向信任者返还增值金额中任意数额的金钱。在该博弈情境中，信任者向被信任者投资的金额(下文统称投资信任水平)即信任，被信任者的返还金额即被信任者的可信度(trustworthiness)。

信任博弈的运用非常广泛(Espín et al., 2016; Bellucci et al., 2019; Javor et al., 2015)，但其内部效度和生态效度遭受了诸多质疑(Karlan, 2005; Thielmann & Hilbig, 2015)。有研究者甚至怀疑信任博弈中的投资行为根本无法测量信任(Ermisch & Gambetta, 2006; Sapienza et al.,

收稿日期: 2020-02-13

* 四川省心理健康教育研究中心一般项目(XLJKJY2016B)

通讯作者：刘昌, E-mail: liuchang@njnu.edu.cn

2013)。鉴于信任博弈对信任研究的重要作用，综述与之相关的主要争议，分析和评估争议的合理性，提出未来研究的展望就具有重要意义。

当前，信任博弈面临的主要争议包括两个方面。第一，信任博弈范式自 Berg 等人(1995)创立以来存在诸多变动，但范式变动的合理性尚存疑问(Aksoy et al., 2018; Lenton & Mosley, 2011)。研究发现投资信任水平在范式变动中可能受到社会偏好(social preference)和风险偏好(risk preference)的影响(Aksoy et al., 2018; Derks et al., 2014; Evans & Krueger, 2011; Thielmann & Hilbig, 2015)，从而降低信任博弈的内部效度。第二，信任博弈中投资信任水平与基于量表法测量的信任(survey based trust)(下文统称量表信任)之间的相关水平较低(Ashraf et al., 2006; Etang et al., 2012; Holm & Nystedt, 2008)。有研究者据此质疑信任博弈的效度并不可靠(Brülhart & Usunier, 2012; Karlan, 2005)。下文将首先从这两个角度梳理和分析信任博弈的主要争议，最后在此基础上提出未来研究的展望。

2 信任博弈的范式变动

研究者基于 Berg 等人(1995)的经典范式，主要在初始金额的对等性、投资金额的连续性、可供投资金额数量以及投资增值率等方面进行变动(Carragher et al., 2018; Johnson & Mislin, 2011; Woolley & Fishbac, 2017)。

2.1 初始金额的对等性

经典信任博弈赋予信任者和被信任者同等数量的初始金额。有研究者出于研究成本的考量，不再给被信任者分配初始金额(Schlösser et al., 2015; Woolley & Fishbac, 2017)，随之带来信任者和被信任者初始收益不对等的问题。初始收益的不对等状态，可能会诱发信任者的内疚感和痛苦感，从而基于不平等厌恶向被信任者投资更多，而非出于信任意向(Aksoy et al., 2018)。与此相对，Evans 和 Revelle(2008)指出，信任博弈中不对等的初始金额设置还可能抑制信任者的投资水平，因为初始金额分配的不对等相对提升了信任者的投资能力，此时信任者就可能出于自利目的降低投资信任水平。Johnson 和 Mislin(2011)一项囊括 162 项研究的元分析支持了 Evans 和 Revelle(2008)的推测。他们发现与初始收益设置对等相比，当初始收益不对等时，信任者的投资数额显著减少。此外，初始金额对等性对投资信任水平的影响可能还受到信任者特质等变量的调节。例如，Corgnet 等人(2016)发现，更厌恶初始收益分配不对等的个体表现出更高水平的投资信任。又如，Hong 和 Bohnet (2007)的研究显示，初始收益分配的不对等状态只会抑制美国人口中较低地位群体——女性、少数民族、年轻人以及非

新教徒的投资信任水平。这意味着，信任者可能基于其社会地位和公平信念等因素，差异化地解释初始收益分配的不对等状态，从而表现出或高或低的投资信任水平。

2.2 投资金额的连续性

经典信任博弈的投资金额为连续(Castilla, 2015; Gereke et al., 2018)，即被试可在投资金额中任选一个数字投资。投资金额的连续性设置运用广泛，但也逐渐暴露了弊端。研究发现，被试在信任博弈中倾向给出不高不低的“适中”投资(Bellucci et al., 2019; McAuliffe et al., 2019)，这种投资倾向可能与信任无关，实际反映了个体的利他或公平等动机(Ashraf et al., 2006; Rudnicki et al., 2019)。为规避此问题，有研究者开发出只能在“投资所有金额”或“完全不投资”中迫选的二元信任博弈范式(binary trust game) (Carragher et al., 2018; Evans & Krueger, 2011)。

Evans 和 Krueger (2011)开发的序列信任博弈范式(sequential trust game)是其中代表。范式包括信任者和被信任者两个角色，游戏收益有 4 个水平，按照 T>R>P>S 排序。游戏开始后，信任者需在信任或维持现状中做出选择。在维持现状的条件下，游戏直接结束，信任者和被信任者都得到较低收益水平 P。在信任条件下，游戏进入第二阶段，被信任者需在互惠和背叛之间做出选择。被信任者选择互惠，则每个玩家都得到较高收益水平 R。如果选择背叛，则信任者得到最低水平收益 S，被信任者得到最高水平收益 T。根据游戏规则和收益矩阵，在信任者方面，选择信任可能会比维持现状获得更高的收益(R>P)，但也可能会因遭到背叛而获益更少(S<P)。因此，信任者需评估是否值得承担背叛的风险，继而做出信任或维持现状的决定。研究者推测，二元信任博弈范式能避免投资信任行为被过分解读为社会偏好(Carragher et al., 2018; Yamagishi et al., 2015)。

2.3 可供投资金额数量与投资增值率

研究者给予被试的可供投资金额数量存在差异。欧美研究的金额设置大多在 5~15 美元区间(Bellucci et al., 2019; Javor et al., 2015; Müller & Schwieren, 2019)，少数在 4 美元以下(Woolley & Fishbach, 2017)或 20 美元以上(Drążkowski et al., 2017)。我国研究多在 5~20 元人民币区间(窦凯等人, 2018; 付超等人, 2018; 刘国芳等人, 2017)。本文认为，可供投资金额数量的设置与研究背景中的地域特性、经济特征等外部环境因素有关，不同地域的研究存在一定差异是合理的。问题在于，迄今研究对可供投资金额数量的设置较为随意，缺乏合理性评估。这可能导致投资信任水平显著受到可供投资金额数量设置的影响，继而降低信任博弈的内部效度。研究发现，当信任者可供投资金额越多，其投资面临的潜在损失越大，投资金额

就越低(Evans & Krueger, 2011)。Johansson-Stenman 等人(2013)的研究也表明，当可供投资金额增加时，个体投资信任的水平出现显著下滑。Johnson 和 Mislin(2011)的元分析同样发现，可供投资金额的多寡能显著预测投资信任水平。

投资增值率，即信任者向被信任者投资金额的增值水平。Berg 等人(1995)的经典范式赋予投资金额三倍增值。有研究者将投资增值率改为二倍(Gereke et al., 2018; Woolley & Fishbach, 2017)或将投资增值率设置为低于三倍增值的某一金额，如将 2 美元增值为 5 美元(Levine & Schweitzer, 2015)。投资增值率与投资信任水平之间存在密切关联。当投资增值率提高，被试投资的理论收益水平随之提升，此时被试更愿意承担风险，做出投资信任行为。反之，被试的投资信任动机则可能被抑制(Evans & Krueger, 2011; Johnson & Mislin, 2011)。例如，Lenton 和 Mosley(2011)考察了投资增值率与投资信任水平的关系。结果发现，投资信任水平随着增值倍率的增加而显著提高。在投资增值率分别为二倍、三倍以及四倍的情境中，被试投资金额占初始金额比例分别为 30%、48%、53%。

2.4 范式变动对测量效度的影响

信任博弈的范式变动为拓展信任博弈的应用范围，提升研究效度提供了新思路(如，二元信任博弈)。但是，部分范式变动缺乏充分的证据支持，尚存一定随意性。例如，研究者降低可供投资金额、取消被信任者初始金额等举措虽然节约了实验成本，但也引发诸多争议。此类争议的核心是范式变动与测量效度的关系问题。信任博弈作为信任测量范式，其信任行为应由信任意向驱动。但当范式变动后，信任者的投资行为被其它变量(如，社会偏好、风险偏好等)赋予更强的解释力，降低了信任博弈的内部效度(Bicchieri et al., 2011)。因此，研究者针对范式变动的争议，本质上反映了社会偏好和风险偏好对投资信任水平的影响问题。有大量研究发现，投资信任水平与社会偏好(Aksoy et al., 2018; Derks et al., 2014; Johansson-Stenman et al., 2013; Sapienza et al., 2013)和风险偏好(Fetchenhauer & Dunning, 2012; Thielmann & Hilbig, 2015)的关系密切，有研究者正基于此怀疑信任博弈无法充分表征信任(Ermisch & Gambetta, 2006; Sapienza et al., 2013)。下文将从社会偏好和风险偏好两个方面分析它们对信任博弈的影响。

2.4.1 社会偏好的影响

社会偏好主要与投资金额对等性和连续性的设置争议相关。根据个体合作性影响信任行为（如通过社会投射）(social projection)的观点，以利他主义(altruism)、互惠(reciprocity)以及不平等厌恶(inequity aversion)为代表的社会偏好被认为是信任的重要预测变量(Aksoy et

al., 2018; Derks et al., 2014; Johansson-Stenman et al., 2013)。例如，Yamagishi 等人(2015)发现，亲社会偏好与投资信任水平呈显著正相关。Naef 和 Schupp(2009)的研究显示，越常参加志愿者服务的被试，其投资信任水平越高。

在社会偏好影响投资信任的机制方面，强互惠理论(strong reciprocity)指出，人们能从互惠行为的过程本身获得效用满足和收益(Gintis, 2000)。有研究者认为，在信任博弈中，向他人投资的行为内涵包括“给予”他人这一心理过程，个体则能从此过程中获得效能感(Ozyilmaz et al., 2018)。在 Dunn 等人(2008)于《Science》上发表的经典研究中，调查数据和实验结果都表明，与给自己花钱相比，人们给他人花钱会感觉更加幸福。这意味着，投资行为本身就受到亲社会、不平等厌恶等因素的影响。因此，用信任博弈中“给予”他人金钱的行为表征信任，可能混淆信任和社会偏好这两个变量(Espín et al., 2016)。

为了厘清这一问题，研究者开展了一系列验证性工作。虽然社会偏好与投资信任之间存在密切关联的现象屡有发现(Fairley et al., 2016; Rudnicki et al., 2019)，但很多研究结果并不支持社会偏好显著影响投资信任行为的假设(Ashraf et al., 2006; Bellucci et al., 2019; Brülhart & Usunier, 2012; Javor et al., 2015)。例如，在 Brülhart 和 Usunier(2012)的研究中，若依据优势利他主义理论假设(dominant altruism)，被试作为信任者，应更倾向于给初始金额分配较少的被信任者投资更多，而向初始金额较多的被信任者投资更少。结果却显示，被试投资金额与被信任者所持金额多寡并无显著相关。总的来说，社会偏好与投资信任水平的关系尚不明朗，有待更多研究予以说明。不过鉴于社会偏好的潜在作用机制，研究者常在实验中对社会偏好进行施测，从而通过统计分析的手段评估其对研究结果的影响程度(Corgnet et al., 2016; Schlösser et al., 2015)。

2.4.2 风险偏好的影响

风险偏好主要与可供投资金额及投资增值率设置的争议相关。在经济学中，风险偏好是个体在投资选择过程中对待风险的心理反应、态度趋向和投资意愿(Markowitz, 1952)，包括风险规避(risk aversion)、风险寻求(risk proneness)、风险中立(risk neu–tralness)的三维结构(Pratt, 1964)。从传统经济博弈理论的视角出发，个体在信任博弈中的行为类似风险投资(Fehr, 2009)，易受风险规避或风险寻求等因素的影响(Evans & Krueger, 2011; Fairley et al., 2016; Thielmann & Hilbig, 2015)。因此，个体在信任博弈中对他人投资金额的多少，可能表征为个体的风险偏好水平(Castilla, 2015; Müller & Schwieren, 2019)。

一方面，有研究发现，个体对不确定性（即风险规避）和潜在损失(即损失规避)的一般

态度会影响信任他人的意愿(Karlan, 2005; Thielmann & Hilbig, 2015)。Sapienza 等人(2013)的研究表明，个体风险规避水平每减少一个标准差，其投资信任的平均金额就增加约 17%。类似的发现是，个体在风险投资中对背叛的厌恶倾向，导致人们在投资回报率由他人决定而非随机的情况下，更希望承担较低风险(Fairley et al., 2016)。另一方面，个体的风险寻求倾向驱使人们追逐利益，表现出投资信任行为。例如，Chetty 等人 (in press)的研究发现，个体在彩票选择任务中越偏好风险寻求，其在信任博弈中的投资金额就越多。

也有很多研究未发现风险偏好显著影响投资信任水平(Ben-Ner & Halldorsson, 2010; Fetchenhauer & Dunning, 2012)。例如，Ben-Ner 和 Halldorsson(2010)的研究显示，量表和行为水平的风险规避都未与投资信任水平存在显著相关。Dunning 等人(2014)的研究表明，依据被试汇报的风险容忍度水平，仅 20.4%的被试愿意在信任博弈中向他人投资，但实际操作中却有 62.4%的被试投资。这表明风险偏好对信任博弈的影响是有条件的。考虑到风险偏好对投资信任水平可能产生影响，许多研究者在实验中对风险偏好进行施测，从而评估其对研究结果的影响程度(Breuer et al., 2016; Yamagishi et al., 2015)。

2.5 小结

总体而言，社会偏好和风险偏好对投资信任水平存在非普遍性的影响，且影响程度与信任博弈范式的变动存在一定关联。不过信任博弈范式变动对研究效度的影响以及社会偏好和风险偏好在其中发挥的作用均不清晰，有待更多研究予以说明。

虽然社会偏好和风险偏好的影响问题引发了研究者对信任博弈的质疑 (Aksoy et al., 2018; Rudnicki et al., 2019)，但从信任兼具多学科、多研究取向的研究现状出发，投资信任水平与社会偏好和风险偏好存在密切关联具备理论层面的合理性。从信任概念的角度看，一方面，有研究者认为，信任本质上是一种道德观念或社会情景中的基本规范(Bicchieri et al., 2011; Uslaner, 2002; Reiersen, 2019)。例如，Uslaner(2002)提出，信任是一种稳定的价值观，一种假定他人都可信的道德戒律，不因特定的事件或利益计算而轻易发生改变。福山(2001)也认为，社会中的信任是一个群体对一种价值观存在的普遍认同。据此，信任动机内涵包括利他主义等社会偏好，投资信任水平也就极易受到社会偏好的影响。另一方面，依据“经济人”假设，信任行为实质上是一种风险情景中的投资行为。因此，信任博弈中的投资行为必然囊括理性的风险评估过程，随之与风险偏好密切相关。再从博弈理论的角度看，传统博弈理论推测，囚徒困境中的游戏双方可通过互不信任实现“纳什均衡”(马庆国等人, 2009)。然而，个体在信任博弈中更倾向于表达互信，背离利益最大化的行为。这可能意味着，对物

质财富的追求并非人类的全部动机，人类行为同时受到内在动机和外在动机的驱动，重视道德、情感情绪、社会规范等(姜树广，韦倩, 2013)。这种行为博弈论的研究思路，有助于得出新的“纳什均衡解”，来解释行为博弈实验中发现的许多悖论(马庆国等人，2009)。因此，投资信任水平与社会偏好和风险偏好的密切关联，也深化了人们对信任概念的理解。综上所述，社会偏好和风险偏好可能构成投资信任行为的信任意向来源，如果仅依据其与投资信任水平的正向关系就质疑信任博弈的效度是存在偏颇的。不过，这并不意味着信任博弈等同于风险偏好范式或社会偏好范式，其本质上为信任测量范式。所以，更关键的是需要合理地开发并运用信任博弈范式，尽可能避免投资信任行为被直接解读为社会偏好和风险偏好，从而提高研究的内部效度。

3 与量表信任的低相关问题

有研究者在运用信任博弈的同时施测量表信任，为信任博弈提供校标关联效度的证据(Etang et al., 2012; Yamagishi et al., 2015)。然而却有不少研究表明，量表信任与投资信任之间缺乏稳定可靠的相关证据(Ashraf et al., 2006; Holm & Nystedt, 2008)。与此同时，被试的量表信任水平，被发现与其扮演被信任者时返还的金额数量，即可信度存在更为显著的正相关(Karlan, 2005; McEvily et al., 2012)。例如，Karlan(2005)的研究显示，个体对社会的量表信任水平能够显著预测其可信度，但不能显著预测投资信任表现。信任博弈与量表信任的低相关现象一直备受关注，有研究者据此认为信任博弈缺乏效度(Brülhart & Usunier, 2012; Karlan, 2005)。本文认为，量表信任与信任博弈的低相关现象可从态度与行为差异、测量类型差异以及量表信任和信任博弈不足这三个角度进行解释。

3.1 态度与行为差异

在态度与行为差异方面，量表信任大多采用主观报告法，要求被试对某个观点在李克特量表上打分，这反映了个体的态度倾向。因此，量表信任也被解读为信任态度(trust attitude)(Gächter et al., 2004)。与之相对，信任博弈则表征信任行为。态度与行为虽然存在密切关联，但二者之间的关系并不确定(Shirokova et al., 2016)。甚至有研究者认为态度与行为直接相关程度很低，对心理学家产生前所未有的挑战(Cooper & Croyle, 1984)。量表信任与信任博弈的低相关现象，可纳入态度与行为关系的范畴中进行讨论。

研究发现，个体的意图、态度的强弱、人际关系等因素都可能导致态度与行为之间出现不一致的情况(Smith & Louis, 2008)。态度与行为的关系问题也历经理性行为理论(theory of

reasoned action, TRA)到囊括非理性因素的计划行为理论(theory of planed behavior, TPB)的发展过程(Michaelidou & Hassan, 2014)。研究者逐渐认识到，态度是包含情感与认知的复杂结构，行为也并非完全由理性驱动。行为只反映符合理性的信念、态度、意图等，而在一些特殊的情况下，行为是不能用理性的标准去衡量的(张红涛, 王二平, 2007)。例如，研究发现，内隐态度和外显态度与行为的关系不一致，内隐态度与快速、自动化的行为关系更为紧密，而外显态度则与意识控制的行为相关更加密切(Swanson et al., 2001)。这揭示了行为系统与态度系统的复杂关系受到理性与非理性变量的综合影响。与之相仿，信任也由完全理性观、非理性观发展到有限理性观(许科, 刘永芳, 2007)。信任行为与信任者特质、被信任者特质、情境变量等一系列包含理性与非理性在内的因素有关，其与信任态度之间的关系受到众多因素的影响(Bujang et al., 2017)。因此，量表信任中的态度与信任博弈中的行为呈现低相关现象，可能是由二者表征的态度系统与行为系统的不确定关系导致。

3.2 测量类型差异

有研究者认为，量表信任与信任博弈之间缺乏稳定相关，可能源于在具体的研究中，两类信任测量方法表征的信任类型和信任对象往往不一致(Sapienza et al., 2013)。从领域一般性-特殊性(domain general-specific)角度可将信任划分为一般信任(general trust)和特殊信任(specific trust)。一般信任大多表征个体普遍的，较稳定的以及无特定信任对象的信任倾向(trust propensity)或特质(trait)。例如，“一般来说，你觉得大多数人都是可以信任的，还是觉得在与人交往的过程中要越小心越好?”(“Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?”)。世界价值观调查(word value survey)、GSS 等信任调查中都采用了此类经典的一般信任问题。与之相对，特殊信任一般是对某个或某类具体信任对象的特定信任，如医生信任(Kovacs et al., 2019)，父母信任(Kerr et al., 1999)等。

研究者往往运用量表信任表征一般信任(Bibi & Karim, 2017; Paxton & Glanville, 2015)，甚至有研究者把量表信任等同于一般信任 (Boyer et al., 2012; Gylfason & Olafsdottir, 2017)。本文认为，这种将量表信任等同于一般信任的做法虽然并不严谨，但也反映了研究者运用量表法测量一般信任的惯例。与之相对，研究者既运用信任博弈表征一般信任(Dunning et al., 2014; Martini & Torcal, 2019)，也测量特殊信任(Anderl et al., 2018; Verberne et al., 2015)。新近研究表明，投资信任是一般信任和特殊信任的结合体(Hale et al., 2017)。Hale 等人(2017)同时操纵可信度高和可信度低的虚拟人物与被试进行信任博弈，发现被试给两个对象的投资

金额存在显著正相关，展现了一般性的信任特征。不过，当按照先不可信对象后可信对象的顺序向被试呈现时，被试对二者的投资就存在显著差异。这说明信任投资行为受到可信度的影响，体现了特殊信任倾向。因此，量表信任与信任博弈的低相关现象，可能源于研究者易忽视二者表征的信任类型往往并不一致 (Holm & Nystedt, 2008; Karlan, 2005)。若对施测的信任类型进行匹配，如，均为一般信任或某一种特殊信任，二者就可能具备较高水平的正相关。不少研究验证了这个推测(Etang et al., 2012; Kovacs et al., 2019; Martini & Torcal, 2019; McEvily et al., 2012)。例如，Kovacs 等人(2019)同时使用量表信任和信任博弈考察塞内加尔病患对医生的特殊信任，显示二者存在显著正相关。

3.3 量表信任和信任博弈的不足

从研究范式的设置和运用出发，量表信任和信任博弈均尚存不足之处，这也可能导致二者的低相关现象。在量表信任方面，研究者对以量表法为主导的信任测量方式存在多种质疑(Boyer et al., 2012)。首先，由于社会赞许效应的存在，个体主观汇报信任水平时存在虚报可能。例如，研究发现，个体的量表信任水平，往往与其内隐信任或信任行为不一致(Ermisch et al., 2009)。一般来说，个人在量表中的信任水平高于其实际的行为表现(Burns & Conchie, 2008)。其次，量表信任中的被信任者概念往往缺乏清晰界定，被试可能因个体和社会环境的差异对被信任者的含义形成不同解读，从而干扰实验结果(Abascal & Baldassarri, 2015)。例如，Van Hoorn (2015)的研究表明，受个体主义和集体主义文化的不同影响，个体在解读“大多数人是否值得信任”时，对“大多数人”范围的理解，即信任半径(trust radius)存在显著差异。Van Hoorn (2015)的研究发现，个体主义文化的信任半径更广，集体主义文化的信任半径则相对更狭窄。此外，量表信任还存在一系列囿于研究方法的缺点。例如，主观评定法获得的数据可能不够精确，收集数据的方式缺乏规范性，与实验情境相比缺乏经济激励，个体缺乏较强动机表达真实想法等等(Fehr et al., 2003)。

在信任博弈方面，社会偏好和风险偏好可能对投资信任水平产生影响，导致投资信任水平在范式变动或无关变量的影响下与真实信任意向相比偏高或偏低，造成信任博弈与量表信任的低相关现象。Glaeser 等人(2000)的研究发现信任博弈与量表信任的相关不显著。Aksoy 等人(2018)推测，未发现显著关联可能与 Glaeser 等人(2000)采用的改编范式有关。与 Berg 等人(1995)的信任博弈范式相比，Glaeser 等人(2000)取消了被信任者的初始金额，将投资金额由连续变为二元。Aksoy 等人(2018)针对 Glaeser 等人(2000)的改编范式和 Berg 等人(1995)的经典范式开展了一项比较研究，结果发现，依据 Berg 等人(1995)的经典范式，信任博弈

和量表信任存在显著正相关，而 Glaeser 等人(2000)的改编范式则与量表信任相关不显著。

综上所述，信任博弈与量表信任的低相关现象可能与态度和行为差异、测量类型差异以及信任博弈和量表游戏的不足之处有关。因此，如果仅依据二者相关水平较低就质疑信任博弈的效度，显然是值得商榷的。

4 总结与展望

综上所述，作为运用最广泛的信任测量范式之一，信任博弈虽尚存不足之处，但迄今研究成果并未形成充分的证据链否认其有效性，信任博弈面临的主要争议也能从理论层面予以较妥善解释。同时，有大量研究支持了信任博弈的科学性 (Dunning et al., 2014; Corgnet et al., 2016; Javor et al., 2015)。据此，本文认为，在现有研究条件下，信任博弈仍是信任间接测量范式中较适宜的一种。信任博弈面对的主要争议，也正是推进其不断发展的动力。

4.1 验证信任博弈范式变动的科学性

信任博弈范式存在多种变动，一些变动有助于从理论层面规避无关变量对研究结果产生影响。例如，将可供投资金额由连续改为二元，有利于厘清社会偏好对投资信任水平的影响问题(Amdur & Schmick, 2012; Ermisch et al., 2009)。但是，此类变动带来诸多争议。例如，可供投资金额的增加或减少可能诱发不同维度和不同水平的风险偏好，从而影响信任投资水平(Fairley et al., 2016; Thielmann & Hilbig, 2015)。因此，通过大量实证研究探明信任博弈范式变动的科学性就非常必要。不过，现阶段该领域研究较为缺乏。例如，Johnson 和 Mislin(2011)虽然发现被试可供投资金额会显著影响投资信任水平，但在他们的元分析中，较大可供投资金额的研究很少，尚待更多证据支持。又如，信任博弈在国内虽然得到广泛运用，但系统检验信任博弈范式合理性的研究仍然较为匮乏。与之相关的是，不同收入水平和不同文化背景中的被试，同等数量的可供投资金额可能带来差异化的心理影响。因此，借助跨文化、跨地区的重复性、比较性研究说明信任博弈范式变动的科学性，应当成为未来信任博弈范式研究的重点。另外，考虑到社会偏好与风险偏好可能对投资信任水平产生影响，很多研究者已在研究中对其进行施测，从而借助统计手段评估潜在的研究误差(Corgnet et al., 2016; Breuer et al., 2016; Schlösser et al., 2015)。现阶段，从实验设计和数据处理两个环节对信任博弈的关键无关变量进行主动施控，是提升研究效度的有益举措。

4.2 进一步澄清信任博弈与量表信任的低相关问题

研究者可依据上文介绍的三种思路对信任博弈与量表信任的低相关问题进一步澄清。此

外，本文推测，量表信任与信任博弈的低相关现象，或许还与两种范式表征信任类型的固有差异有关。当运用量表法和信任博弈但并不指定明确信任对象时，量表信任和信任博弈均默认施测一般信任(Corgnet et al., 2016; Drążkowski et al., 2017; Gereke et al., 2018; Gylfason & Olafsdottir, 2017; Levine & Schweitzer, 2015; Martini & Torcal, 2019; Paxton & Glanville, 2015; Wang & Murnighan, 2017)。但问题在于，有研究表明，信任博弈无明确投资对象时，个体投资信任行为或许并不表征一般信任。例如，Naef 和 Schupp(2009)在研究中同时施测了陌生人信任量表(soep–trust)、一般信任 GSS 量表以及没有明确投资对象的信任博弈。结果表明，信任博弈仅与陌生人信任量表(soep–trust)存在显著正相关，与 GSS 相关则不显著。类似发现是，Fehr 等人(2003)的研究显示，未指定信任对象的信任博弈与信任对象设定为"陌生人"的量表信任之间存在显著正相关。这些研究结果暗示，当信任博弈不指定明确的信任对象时，其信任类型是针对"陌生人"群体的特殊信任，而非一般信任。一般信任表征的信任对象为"大多数人"，包含了陌生人和熟人，与陌生人信任并不等同。据此推测，若量表信任和信任博弈的信任对象并未指明时，量表信任表征一般信任，而信任博弈指代陌生人信任，则量表信任与信任博弈的低相关问题，或许还与两类测量方法在表征信任内涵方面存在固有差异有关。

进一步推测，不指定对象的信任博弈与陌生人信任的关系更为紧密，或许源自于信任博弈的范式特性。在实验室情境中，被试的信任对象无论为真实或虚拟，在未说明信任对象相关信息的条件下，实际上默认了被信任者是信任者的"陌生人"。此外，信任博弈模拟真实信任问题情境，近似现实情境任务(real-world tasks)，范式中包含了社会情境(social presses)的影响。虽然信任博弈中的社会情境是否会诱发个体针对陌生人的信任感尚缺乏证据，不过有其它领域的研究表明，当涉及问题解决的测量范式为现实情境任务时，社会情境促使个体更多地从负向思考问题。例如，Kapoor 和 Khan(2019)创造力领域的研究表明，当创造力问题包括现实情境，被试会表现出更高水平的负性创造力。个体对社会负性信息具备自动警觉的特性(Pratto & John, 1991)，在人际交互活动中，人们倾向于把陌生人识别为负性因素，即不确定和可能不值得信任的(Feldmanhall et al., 2018)，这或许会令被试在信任博弈中对人际信任的风险因素更为警觉。据此，信任博弈可能唤醒个体更多关注人际互动的负性线索，从而将不确定被信任对象的信任博弈解读为陌生人博弈，进而导致信任博弈与陌生人信任的关系更为紧密。当然，这只是基于当前研究结果的大胆推测，尚待更多研究检验。

4.3 拓展信任博弈的测量维度

[在此处键入]

信任是多维度概念。但如前文所述，信任博弈和量表信任测量的信任类型不一致可能带来二者的低相关问题。与量表信任相比，信任博弈施测的信任维度非常单一，这给信任博弈的应用带来一定的隐患和限制。例如，信任可依据被信任者特征划分为能力信任(ability trust)、正直信任(integrity trust)以及善意信任(benevolence trust) (Mayer et al., 1995)。正直信任易与善意信任混淆，但二者是不同的信任维度，有相对独立的发生机制(Schweitzer et al., 2006)。Levine 和 Schweitzer(2015)的实验对信任博弈隐含的心理结构进行探索，分离出信任博弈中的善意和正直成分。他们发现，相较于诚实但自私的对象，被试更信任欺骗但利他的个体。据此，当前信任博弈范式中个体的信任行为可能主要基于善意层面，很少涉及能力和正直(Levine & Schweitzer, 2015)。然而，人际信任不仅有赖于被信任者善意，能力和正直特质同样重要。因此，信任博弈单一的信任维度表征，不仅可能诱发信任博弈和量表信任的低相关，还制约了信任博弈的测量功效。未来可尝试基于 Mayer 等人(1995)的信任理论，从被信任者特质影响信任维度的思路着手，通过合理调整范式，拓展信任博弈的测量维度。例如，修改信任博弈的问题情境，使被试投资金额能否获益取决于被信任者的知识和能力而非善意，继而表征能力信任。

4.4 提升信任博弈的生态效度

信任博弈模拟了现实生活中人际信任的发生过程，与量表信任相比，其生态效度更具优势。但在实际运用中，博弈情景却较为失真。首先，在现实社会中，人们往往同时面对多个潜在信任对象，而非仅限于囚徒困境式的单对单博弈。为了弥补此类不足，有研究者将经典信任博弈的角色设置更改为一名信任者和两名被信任者，借此模拟现实中多被信任者的复杂情境(张希等人, 2019)。此外，传统信任博弈范式较为抽象，不适用于儿童被试。有研究针对儿童被试调整信任博弈的奖励机制，如将金钱改为糖果，将复杂的奖惩机制表述为简明直白的互动游戏规则，令儿童体验到更真实的信任发生过程(陈丽君等人, 2018)。除此之外，传统信任博弈一般在实验室情境中进行或以书面文字的形式呈现(Gereke et al., 2018; Righetti et al., 2015)，被信任者的信息大多不被告知(Bauer & Freitag, 2018; Wang & Murnighan, 2017)，信任者和被信任者之间一般也没有交流(Drążkowski et al., 2017)。为弥补此类不足，研究者运用虚拟现实技术进一步提升信任博弈的生态效度。Hale 等人(2017)基于信任博弈的基本范式开发了虚拟迷宫(virtual maze)测验。被试佩戴虚拟现实仪器进入程序后，会发现自身处于由多个相似房间构成的走廊中。被试的任务为根据游戏中虚拟人物呈现的信息，选择进入某些房间，最终脱离迷宫。主试通过记录被试是否咨询虚拟人物，是否听

从虚拟人物的意见来综合判定其信任水平。虚拟现实技术拥有良好的拟真性，研究者能借此记录被试更为真实的行为反应。未来可基于此开发更具生态效度的信任博弈测验，进一步丰富信任测量技术。

参考文献

陈丽君，王欣，赵陵波，陈昕，王益文. (2018). 面孔二态性对学前儿童信任行为的影响——来自人格标签的解释. *心理发展与教育*, *34*(5), 3–12.

窦凯，聂衍刚，王玉洁，刘耀中. (2018). 信任还是设防？互动博弈中社会善念对合作行为的促进效应. *心理科学*, *41*(2), 390–396.

福山. (2001). *信任：社会美德与创造经济繁荣*. 海南出版社.

付超，张振，何金洲，黄四林，仇剑鋆，王益文. (2018). 普遍信任博弈决策的动态过程——来自脑电时频分析的证据. *心理学报*, *50*(3), 317–326.

姜树广，韦倩. (2013). 信念与心理博弈:理论、实证与应用. *经济研究*, *48*(6), 141–154.

刘国芳，辛自强，林崇德. (2017). 人际信任中的坏苹果效应及其传递. *心理与行为研究*, *15*(5), 691–696.

马庆国，沈强，李典典，王凯. (2009). 社会神经经济学：社会决策和博弈的神经学基础. *浙江大学学报（人文社会科学版）*, *39*(2), 53–63.

许科，刘永芳. (2007). 有限理性信任观:对理性计算和非理性态度的整合. *心理科学*, *30*(5), 1193–1195.

张红涛，王二平. (2007). 态度与行为关系研究现状及发展趋势. *心理科学进展*, *15*(01), 163–168.

张宁，张雨青，吴坎坎. (2011). 信任的心理和神经生理机制. *心理科学*, *34*(5), 1137–1143.

张希，黄登仕，董占奎. (2019). 社会关系网络中信息不对称对信任行为影响的实验研究. *系统管理学报*. *28*(2), 269–276.

Abascal, M. , & Baldassarri, D. (2015). Love thy neighbor? Ethnoracial diversity and trust reexamined 1. *American Journal of Sociology*, *121*(3), 722–782.

Aksoy, B., Harwell, H., Kovaliukaite, A., & Eckel, C. (2018). Measuring trust: A reinvestigation. *Southern Economic Journal*, *84*(4), 992–1000.

Amdur, D., & Schmick, E. (2012). Does the direct-response method induce guilt aversion in a trust game? *Economics Bulletin*, *33*(1), 687–693.

Anderl, C., Steil, R., Hahn, T., Hitzeroth, P., Reif, A., & Windmann, S. (2018). Reduced reciprocal giving in social anxiety-Evidence from the Trust Game. *Journal of Behavior Therapy and*

Experimental Psychiatry, *59*, 12–18.

Ashraf, N., Bohnet, I., & Plankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, *9*(3), 193–208.

Bauer, P. C., & Freitag, M. (2018). Measuring trust. In E. Uslaner(Eds.), *The Oxford Handbook of Social and Political Trust* (pp.15–36). Oxford Press.

Bellucci, G., Hahn, T., Deshpande, G., & Krueger, F. (2019). Functional connectivity of specific resting–state networks predicts trust and reciprocity in the trust game. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(3), 165–176.

Ben-Ner, A., & Halldorsson, F. (2010). Trusting and trustworthiness: What are they, how to measure them, and what affects them. *Journal of Economic Psychology*, *31*, 64–79.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*(1), 122–142.

Bibi, Z., & Karim, J. (2017). Self–esteem, resilience, and social trust as mediators in the relationship between optimism and job satisfaction: a preliminary analysis of data from european social survey. *Pakistan Journal of Psychological Research*, *32*(1), 155–173.

Bicchieri, C., Xiao, E., & Muldoon, R. (2011). Trustworthiness is a social norm, but trusting is not. Politics, *Philosophy & Economics*, *10*(2), 170–187.

Boyer, P., Lienard, P., & Xu, J. (2012). Cultural differences in investing in others and in the future: Why measuring trust is not enough. *Plos One*, *7*(7), 1–5.

Breuer, W., Helduser, C., & Schade, P. (2016). Breaking the rules: Anticipation of norm violation in a binary-choice trust game. *Economics Letters*, *146*, 123–125.

Brülhart, M., & Usunier, J. C. (2012). Does the trust game measure trust? *Economics Letters*, *115*(1), 20–23.

Bujang, A., Suki, N. M., & Suki, N. M. (2017). Structural relationship between women consumers' attitude and actual behavior in mobile retailing. *Advanced Science Letters*, *23*(9), 8160–8163.

Burns, C., & Conchie, S. M. (2008). Measuring implicit trust and automatic attitude activation. In F. Lyon., G. Mšllering, & M. N. Saunders(Eds.) (pp.239–248), *Handbook of research methods on trust*. Edward Elgar Publishing.

Carragher, D. J., Thomas, N. A., & Nicholls, M. E. (2018). Is trustworthiness lateralized in the face? Evidence from a trust game. *Laterality: Asymmetries of Body, Brain and Cognition*, *23*(1), 1–19.

Castilla, C. (2015). Trust and reciprocity between spouses in India. *American Economic Review*, *105*(5), 621–624.

Chetty, R., Hofmeyr, A., Kincaid, H., & Monroe, B. (in press). The trust game does not (only) measure trust: The risk-trust confound revisited. *Journal of Behavioral and Experimental Economics*, 101520.

Cooper, J., & Croyle, R. T. (1984). Attitudes and attitude change. *Annual Review of Psychology*, *35*(1), 395–426.

Corgnet, B., Espín, A. M., Hernán-González, R., Kujal, P., & Rassenti, S. (2016). To trust, or not to trust: Cognitive reflection in trust games. *Journal of Behavioral and Experimental Economics*, *64*, 20–27.

Derks, J., Lee, N. C., & Krabbendam, L. (2014). Adolescent trust and trustworthiness: Role of gender and social value orientation. *Journal of Adolescence*, *37*(8), 1379–1386.

Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, *2*, 265–279.

Drążkowski, D., Kaczmarek, L. D., & Kashdan, T. B. (2017). Gratitude pays: A weekly gratitude intervention influences monetary decisions, physiological responses, and emotional experiences during a trust-related social interaction. *Personality and Individual Differences*, *110*, 148–153.

Dunn, E. W., Aknin, L. B., & Norton, M. I. (2008). Spending money on others promotes happiness. *Science*, *319*(5870), 1687–1688.

Dunning, D., Anderson, J. E., Schlösser, T., Ehlebracht, D., & Fetchenhauer, D. (2014). Trust at zero acquaintance: More a matter of respect than expectation of reward. *Journal of Personality and Social Psychology*, *107*(1), 122–141.

Ermisch, J., & Gambetta, D. (2006). People's trust: The design of a survey-based experiment. IZA Discussion Papers 2216. *Institute for the Study of Labor*.

Ermisch, J., Gambetta, D., Laurie, H., Siedler, T., & Noah Uhrig, S. C. (2009). Measuring people's trust. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(4), 749–769.

Espín, A. M., Exadaktylos, F., & Neyse, L. (2016). Heterogeneous motives in the trust game: A

tale of two roles. *Frontiers in Psychology*, *7*(1090), 728.

Etang, A., Fielding, D., & Knowles, S. (2012). Are survey measures of trust correlated with experimental trust? Evidence from Cameroon. *The Journal of Development Studies*, *48*(12), 1–15.

Evans, A. M., & Krueger, J. I. (2011). Elements of trust: Risk and perspective taking. *Journal of Experimental Social Psychology*, *47*(1), 171–177.

Evans, A. M., & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, *42*(6), 1585–1593.

Fairley, K., Sanfey, A. G., Vyrastekova, J., & Weitzel, U. (2016). Trust and risk revisited. *Journal of Economic Psychology*, *57*, 74–85.

Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, *7*, 235–266.

Fehr, E., Fischbacher, U., Rosenbladt, B., Schupp, J., & Wagner, G. (2003). A nation–wide laboratory-Examining trust and trustworthiness by integrating experiments in representative surveys. *Institute for Empirical Research in Economics, University of Zürich, Working Paper.*

FeldmanHall, O., Dunsmoor, J. E., Tompary, A., Hunter, L. E., Todorov, A., & Phelps, E. A. (2018). Stimulus generalization as a mechanism for learning to trust. *Proceedings of the National Academy of Sciences*, *115*(7), 1690–1697.

Fetchenhauer, D., & Dunning, D. (2012). Betrayal aversion versus principled trustfulness-how to explain risk avoidance and risky choices in trust games. *Journal of Economic Behavior and Organization*, *81*(2), 534–541.

Gächter, S., Herrmann, B., & Thöni, C. (2004). Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. *Journal of Economic Behavior & Organization*, *55*(4), 505–531.

Gereke, J., Schaub, M., & Baldassarri, D. (2018). Ethnic diversity, poverty and social trust in Germany: Evidence from a behavioral measure of trust. *Plos One*, *13*(7), 1–15.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206*(2), 169–179.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *The*

Quarterly Journal of Economics, *115*(3), 811–846.

Gylfason, H. F., & Olafsdottir, K. (2017). Does Gneezy's cheap talk game measure trust? *Journal of Behavioral and Experimental Economics*, *67*, 143–148.

Hale, J., Payne, M. E., Taylor, K. M., Paoletti, D., & Hamilton, A. F. D. C. (2017). The virtual maze: A behavioural tool for measuring trust. *The Quarterly Journal of Experimental Psychology*, *71*(4), 989–1008.

Holm, H., & Nystedt, P. (2008). Trust in surveys and games-a methodological contribution on the influence of money and location. *Journal of Economic Psychology*, *29*(4), 522–542.

Hong, K., & I. Bohnet (2007). Status and distrust: The relevance of inequality and betrayal aversion, *Journal of Economic Psychology*, *28*, 197–213.

Javor, A., Riedl, R., Kirchmayr, M., Reichenberger, M., & Ransmayr, G. (2015). Trust behavior in Parkinson’s disease: Results of a trust game experiment. *BMC Neurology*, *15*(1), 126.

Johansson–Stenman, O., Mahmud, M., & Martinsson, P. (2013). Trust, trust games and stated trust: Evidence from rural Bangladesh. *Journal of Economic Behavior & Organization*, *95*, 286–298.

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*(5), 865–889.

Kapoor, H., & Khan, A. (2019). Creativity in context: Presses and task effects in negative creativity. *Psychology of Aesthetics, Creativity, and the Arts. 13*(3), 314–321.

Karlan, D. S. (2005). Using experimental economics to measure social capital and predict financial decisions. *American Economic Review*, *95*(5), 1688–1699.

Kerr, M., Stattin, H., & Trost, K. (1999). To know you is to trust you: Parents' trust is rooted in child disclosure of information. *Journal of Adolescence*, *22*(6), 737–752.

Kovacs, R. J., Lagarde, M., & Cairns, J. (2019). Measuring patient trust: Comparing measures from a survey and an economic experiment. *Health Economics*, *28*(5), 641–652.

Lenton, P., & Mosley, P. (2011). Incentivising trust. *Journal of Economic Psychology*, *32*(5), 890–897.

Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, *126*, 88–106.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, *7*(1), 77–91.

Martini, S., & Torcal, M. (2019). Trust across political conflicts: Evidence from a survey experiment in divided societies. *Party Politics*, *25*(2), 126–139.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734.

McAuliffe, W. H., Forster, D. E., Pedersen, E. J., & McCullough, M. E. (2019). Does cooperation in the laboratory reflect the operation of a broad trait? *European Journal of Personality*, *33*(1), 89–103.

McEvily, B., Radzevick, J. R., & Weber, R. A. (2012). Whom do you distrust and how much does it cost? An experiment on the measurement of trust. *Games and Economic Behavior*, *74*(1), 285–298.

Michaelidou, N., & Hassan, L. (2014). New advances in attitude and behavioural decision-making models. *Journal of Marketing Management*, *30*(5-6), 519–528.

Müller, J., & Schwieren, C. (2019). Big five personality factors in the trust game. *Journal of Business Economics*, *90*, 37–55.

Naef, M., & Schupp, J. (2009). Measuring trust: experiments and surveys in contrast and combination. *Working Paper, Royal Holloway College, London*.

Ozyilmaz, A., Erdogan, B., & Karaeminogullari, A. (2018). Trust in organization as a moderator of the relationship between self-efficacy and workplace outcomes: A social cognitive theory-based examination. *Journal of Occupational and Organizational Psychology*, *91*(1), 181–204.

Paxton, P., & Glanville, J. L. (2015). Is trust rigid or malleable? A laboratory experiment. *Social Psychology Quarterly*, *78*(2), 194–204.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica, 32,* 122–136.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380–391.

Reiersen, J. (2019). Drivers of trust and trustworthiness. *International Journal of Social Economics*, *46*(1), 2–17.

Righetti, F., Balliet, D., Visserman, M., & Hofmann, W. (2015). Trust and the suppression of emotions during sacrifice in close relationships. *Social Cognition*, *33*(5), 505–519.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, *35*(4), 651–665.

Rudnicki, K., De Backer, C. J., & Declerck, C. (2019). The effects of celebrity gossip on trust are moderated by prosociality of the gossipers. *Personality and Individual Differences*, *143*, 42–46.

Sapienza, P., Toldra-Simats, A., & Zingales, L. (2013). Understanding trust. *The Economic Journal*, *123*(573), 1313–1332.

Schlösser, T., Mensching, O., Dunning, D., & Fetchenhauer, D. (2015). Trust and rationality: Shifting normative analyses of risks involving other people versus nature. *Social Cognition*, *33*(5), 459–482.

Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes*, *101*(1), 1–19.

Shirokova, G., Osiyevskyy, O., & Bogatyreva, K. (2016). Exploring the intention-behavior link in student entrepreneurship: Moderating effects of individual and environmental characteristics. *European Management Journal*, *34*(4), 386–399.

Smith, J. R., & Louis, W. R. (2008). Do as we say and as we do: The interplay of descriptive and injunctive group norms in the attitude-behaviour relationship. *British Journal of Social Psychology, 47*(4), 647–666.

Swanson, J. E., Swanson, E., & Greenwald, A. G. (2001). Using the implicit association test to investigate attitude-behaviour consistency for stigmatised behaviour. *Cognition and Emotion*, *15*(2), 207–230.

Thielmann, I., & Hilbig, B. E. (2015). Trust: An integrative review from a person–situation perspective. *Review of General Psychology*, *19*(3), 249–277.

Uslaner, E. M. (2002). The Moral Foundations of Trust. *Cambridge University Press*.

Van Hoorn, A. (2015). Individualist-collectivist culture and trust radius: A multilevel approach. *Journal of Cross-Cultural Psychology*, *46*(2), 269–276.

Verberne, F. M., Ham, J., & Midden, C. J. (2015). Trusting a virtual driver that looks, acts, and thinks like you. *Human Factors*, *57*(5), 895–909.

Wang, L., & Murnighan, J. K. (2017). The dynamics of punishment and trust. *Journal of Applied Psychology*, *102*(10), 1385–1402.

Woolley, K., & Fishbach, A. (2017). A recipe for friendship: Similar food consumption promotes trust and cooperation. *Journal of Consumer Psychology*, *27*(1), 1–10.

Yamagishi, T., Akutsu, S., Cho, K., Inoue, Y., Li, Y., & Matsumoto, Y. (2015). Two-component model of general trust: Predicting behavioral trust from attitudinal trust. *Social Cognition*, *33*(5), 436–458.

Can trust game measure trust? Controversy and Outlook

GONG Zhe[1]; TANG Yujie[1]; LIU Chang[2]

([1]*School of Psychology, Sichuan Normal University, Chengdu* 610068, *China*)

([2] *School of Psychology, Nanjing Normal University, Nanjing* 210097, *China)*

Abstract: Trust games are widely used as a classic paradigm of trust measurement. However, the accuracy of the results of the trust game had been questioned by scholars in the field all the time, making precise trust measurement an important topic for exploration. The controversy mainly includes the following two aspects. On the one hand, the appropriateness of the trust game paradigm has yet to be verified. The controversy over paradigm changes further reveals that social preferences and risk preferences might have an impact on the level of investment trust, thereby reducing the internal validity of the trust game. On the other hand, the correlation between the level of the investment trust in trust games and the trust measured by the survey is very low, which can be explained by differences in attitudes and behaviours, differences in measurement types, and the limitations of trust measured by the survey and trust game. Although trust game face with some controversy, in general, it is still a suitable method for trust measuring. Future research should focus on the following points: Verifying the scientificity of the changes in the paradigm of trust game; Clarifying the low correlation between trust game and trust measured by survey; Expanding the measurement dimensions of trust game; Improving the ecological validity of trust game.

Key words: trust; trust game; social preference; risk preference